

Abstract

A technique for adaptation of a speech recognizing system across multiple remote communication sessions with a speaker. The speaker can be a telephone caller. An acoustic model is utilized for recognizing the speaker's speech. Upon initiation of a first remote session with the speaker, the acoustic model is speaker-independent. During the first session, the speaker is uniquely identified and speech samples are obtained from the speaker. In the preferred embodiment, the samples are obtained without requiring the speaker to engage in a training session. The acoustic model is then modified based upon the samples thereby forming a modified model. The model can be modified during the session or after the session is terminated. Upon termination of the session, the modified model is then stored in association with an identification of the speaker. During a subsequent remote session, the speaker is identified and, then, the modified acoustic model is utilized to recognize the speaker's speech. Additional speech samples are obtained during the subsequent session and, then, utilized to further modify the acoustic model. In this manner, an acoustic model utilized for recognizing the speech of a particular speaker is cumulatively modified according to speech samples obtained during multiple sessions with the speaker. As a result, the accuracy of the speech recognizing system improves for the speaker even when the speaker only engages in relatively short remote sessions.